Electronic Journal of Biotechnology 18 (2015) 61–67

HOSTED BY

ELSEVIER

Contents lists available at ScienceDirect

Electronic Journal of Biotechnology

Identification of groundnut (*Arachis hypogaea*) SSR markers suitable for multiple resistance traits QTL mapping in African germplasm

Busisiwe T. Ncube Kanyika [a], Davies Lungu [a], Alice M. Mweetwa [b], Evans Kaimoyo [c], Vincent M. Njung'e [d], Emmanuel S. Monyo [d], Moses Siambi [d], Guohao He [e], Channapata S. Prakash [e], Yongli Zhao [e], Santie M. de Villiers [d,f,*]

[a] *Plant Science Department, School of Agricultural Sciences, University of Zambia, Box 32379, Lusaka, Zambia*
[b] *Soil Science Department, School of Agricultural Sciences, University of Zambia, Box 32379, Lusaka, Zambia*
[c] *Biological Sciences Department, School of Agricultural Sciences, University of Zambia, Box 32379, Lusaka, Zambia*
[d] *ICRISAT—Nairobi, PO Box 36093 00623, Nairobi, Kenya*
[e] *Department of Agricultural and Environmental Sciences, College of Agriculture, Environment and Nutrition Studies, Tuskegee University, AL, USA*
[f] *Department of Chemistry and Biochemistry, Pwani University, PO Box 195, 10801 Kilifi, Kenya*

ARTICLE INFO

Article history:
Received 26 May 2014
Accepted 1 October 2014
Available online 10 November 2014

Keywords:
Arachis
African varieties
Disease
Polymorphism
SSR

ABSTRACT

Background: This study aimed to identify and select informative Simple Sequence Repeat (SSR) markers that may be linked to resistance to important groundnut diseases such as Early Leaf Spot, Groundnut Rosette Disease, rust and aflatoxin contamination. To this end, 799 markers were screened across 16 farmer preferred and other cultivated African groundnut varieties that are routinely used in groundnut improvement, some with known resistance traits.

Results: The SSR markers amplified 817 loci and were graded on a scale of 1 to 4 according to successful amplification and ease of scoring of amplified alleles. Of these, 376 markers exhibited Polymorphic Information Content (PIC) values ranging from 0.06 to 0.86, with 1476 alleles detected at an average of 3.7 alleles per locus. The remaining 423 markers were either monomorphic or did not work well. The best performing polymorphic markers were subsequently used to construct a dissimilarity matrix that indicated the relatedness of the varieties in order to aid selection of appropriately diverse parents for groundnut improvement. The closest related varieties were MGV5 and ICGV-SM 90704 and most distant were Chalimbana and 47–10. The mean dissimilarity value was 0.51, ranging from 0.34 to 0.66.

Discussion: Of the 376 informative markers identified in this study, 139 (37%) have previously been mapped to the *Arachis* genome and can now be employed in Quantitative Trait Loci (QTL) mapping and the additional 237 markers identified can be used to improve the efficiency of introgression of resistance to multiple important biotic constraints into farmer-preferred varieties of Sub-Saharan Africa.

1. Introduction

Cultivated groundnut or peanut (*Arachis hypogaea* L.) is a cleistogamous allotetraploid leguminous annual crop with a genome of 2891 Mbp [1]. In Africa, where undernourishment from 2007–2008 increased by 10% with an increase in the price of nutritious foods, groundnut is an important cash crop, an affordable source of edible oil rich in omega-3 fatty acids, protein and vitamin E and its stover provides nutritious fodder for livestock [2,3,4]. Yield per hectare in Eastern and South Central Africa averages 1604 kg/ha, which is low compared to the 3393 kg/ha and 3801 kg/ha routinely harvested in China and the United States of America, respectively [4]. A major constraint to achieving the yield potential of groundnuts in Eastern and Southern Africa has been the prevalence of viral Groundnut Rosette disease (GRD), fungal rust and Early Leaf Spot (ELS) diseases [5]. *Aspergillus flavus/parasiticus* is also an important fungus that attacks groundnut post-harvest since consumption of aflatoxins can result in death [6] and its presence inevitably lowers yield quality.

The high cost of chemicals limits control of groundnut diseases in Africa and its use depends on ideal weather conditions, cultural practices and good application skills [7,8,9,10]. Biological control studies with mycoparasites [11] and *Bacillus cereus* [12] have been successful but limited to controlled environments.

Groundnuts exhibit low outcrossing rates ranging from 0 to 8% [13,14,15] and innate disease resistance is seldom attained through natural outcrossing. Historically, introgression of existing resistance

* Corresponding author.
E-mail address: s.devilliers@pu.ac.ke (S.M. de Villiers).
Peer review under responsibility of Pontificia Universidad Católica de Valparaíso.

http://dx.doi.org/10.1016/j.ejbt.2014.10.004

Table 1
African *Arachis* germplasm used in this study grouped according to their attributes of disease tolerance/resistance, productivity and quality traits and farmer preference.

Category	Genotype	Essential traits		Country of cultivation
		Disease resistance/susceptibility	Other agronomic traits	
Disease resistance/ tolerance	ICGV-SM 95342	LLS and rust resistant	–	Malawi
	ICGV 94114	Rust resistant (Good parent for resistance breeding)	–	Malawi
	ICG 12991	*Aphis* sp. resistance (GRD)	Spanish, short duration, drought-tolerant	India, Malawi, Mozambique, Uganda, Zambia
	ICGV-SM 90704	GRV resistant, *Aphis* sp. susceptible	Virginia bunch type, high-yielding, medium–duration, difficult to shell	Malawi, Uganda, Mozambique, Zambia
	ICG 7878	LLS resistant, ELS tolerant	Virginia bunch type, amenable to technology, large seeds	–
	ICGV 95714	ELS resistant (Good parent for resistance breeding)	Short duration	–
	55–437	Aflatoxin tolerant	Drought resistant, high oil content	West Africa
High yield and other quality traits	FLEUR II	ELS and aflatoxin susceptible	Non-dormant	–
	CG 7 (MGV 4)	GRD, ELS, rust susceptible	Drought tolerant, good taste, short cooking time, uniform kernels, high oil content	Malawi, Zambia
	MGV 5		Virginia runner type, confectionery, high oil content, roasts well, attractive tan-colored kernels	Zambia
	Chalimbana	GRD, ELS and rust susceptible	Virginia runner type, large seeds, high oil content, easy shelling, good taste, pre-harvest dormancy	Malawi, Zambia
Farmer preferred traits	ICGV-SM 99557		High-yielding	Malawi
	Pendo		High-yielding, large seeds	Tanzania
	ICGV 86124		Spanish, early-maturing, high-yielding	Senegal, Mali.
	47–10	Resistance to *Phythium* sp.	–	–
	JL 24 (Luena)	GRD, ELS, rust susceptible	Spanish, early-maturing, high-yielding, drought tolerant, non-dormant	India, Malawi, Mali, Philippines, DR Congo, Zambia, South Africa, Zimbabwe

and other farmer preferred traits is accomplished only through artificial hybridization in targeted breeding from, for example, diploid wild relatives of groundnut with known abiotic and biotic stress resistance and/or tolerance [5]. In general, inheritance of disease resistance has been governed by quantitative recessive genes with low heritability that are controlled by epistatic effects and the environment [9]. The narrow genetic base of cultivated groundnut and variation in ploidy levels further limits introgression of resistance traits by interspecific hybridization [2].

Detection of polymorphic molecular markers associated with genes governing disease and insect resistance has progressed rapidly over the past two decades. This accelerated the development of cultivar resistance breeding programs for enhanced yield and grain quality [16,17,18]. SSR markers are preferred due to their co-dominance, simplicity, high polymorphism, repeatability, multi-allelic nature and transferability within the genus *Arachis* and significant polymorphism has been identified in novel Simple Sequence Repeat (SSRs) by He et al. [19]. These markers have enhanced phylogenetic studies of the *Arachis* species, for pre-breeding parent determination and integration of SSR based maps in both diploid and tetraploid species [20,21,22], comprehensive Quantitative Trait Loci (QTL) analysis for linkage to disease and pest resistance [23,24,25], comparative

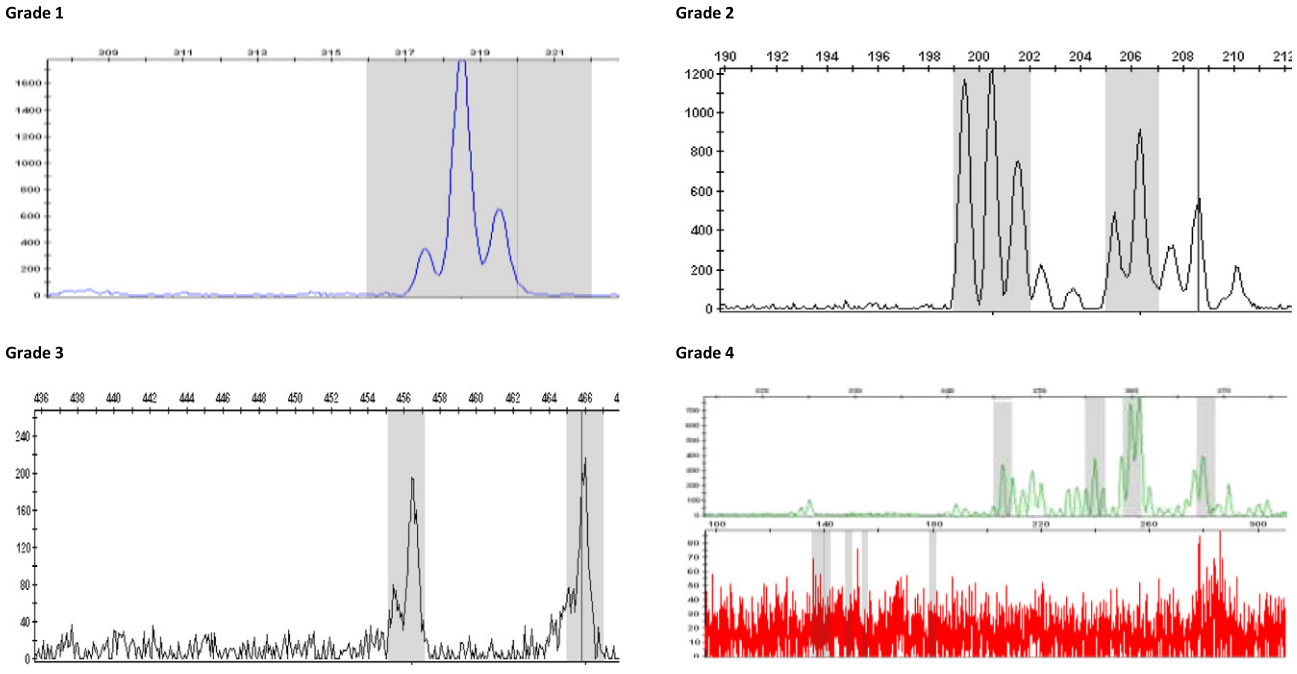

Fig. 1. SSR fragment analysis images showing examples of the different allele grades allocated according to ease of scoring.

Table 2
Dissimilarity matrix of 16 *Arachis* sp. Genotypes. Appropriate disease resistance/tolerance pair wise comparisons between varieties (>0.532) are highlighted for ELS (orange), GRD (red), GRD-aphid (green), rust (blue) and aflatoxin (pink).

Genotype	ICG 7878	ICG 12991	55–437	ICGV 86124	ICGV-SM 90704	ICGV 94114	ICGV 95342	ICGV-SM 95714	ICGV–SM 99557	47-10	CG7	Chalimbana	FLEUR-II	JL24	MGV 5
ICG 12991	0.458														
55–437	0.508	0.407													
ICGV 86124	0.582	0.506	0.407												
ICGV-SM 90704	0.479	0.468	0.546	0.547											
ICGV 94114	0.538	0.458	0.441	0.496	0.542										
ICGV-SM 95342	0.507	0.572	0.552	0.520	0.550	0.544									
ICGV-SM 95714	0.567	0.532	0.491	0.514	0.519	0.543	0.548								
ICGV-SM 99557	0.532	0.452	0.442	0.488	0.499	0.427	0.549	0.499							
47–10	0.607	0.504	0.383	0.468	0.591	0.509	0.571	0.566	0.511						
CG7	0.513	0.479	0.579	0.551	0.404	0.499	0.537	0.522	0.446	0.611					
Chalimbana	0.409	0.483	0.577	0.594	0.400	0.566	0.534	0.512	0.527	0.662	0.439				
FLEUR–II	0.570	0.526	0.394	0.454	0.536	0.522	0.560	0.503	0.487	0.493	0.593	0.543			
JL24	0.597	0.532	0.419	0.412	0.563	0.542	0.567	0.567	0.525	0.419	0.615	0.580	0.425		
MGV 5	0.471	0.485	0.547	0.567	0.347	0.549	0.535	0.523	0.518	0.651	0.438	0.310	0.533	0.589	
PENDO	0.532	0.471	0.475	0.419	0.540	0.509	0.594	0.528	0.452	0.526	0.511	0.517	0.447	0.370	0.563

mapping studies [26,27] and as a basis for identification of candidate genome regions controlling rust and LLS resistance [28,29]. Wang et al. [30] constructed a genetic linkage map from SSR derived bacterial artificial chromosome end sequences, facilitating the identification of markers linked to resistance gene homologs and map-based cloning. Even markers with low polymorphism enhanced the total available SSRs in wild species for transfer of target traits and should not be disregarded [31].

This study was undertaken to identify and select informative SSR markers that may be linked to resistance to ELS, GRD, rust and aflatoxin contamination across 16 varieties of farmer-preferred and other cultivated African groundnut varieties that are routinely used in groundnut improvement in order to aid the identification of suitable parents for mapping populations or marker-assisted introgression and to select a subset of SSR markers that are evenly spread across the groundnut genome for future resistance QTL mapping.

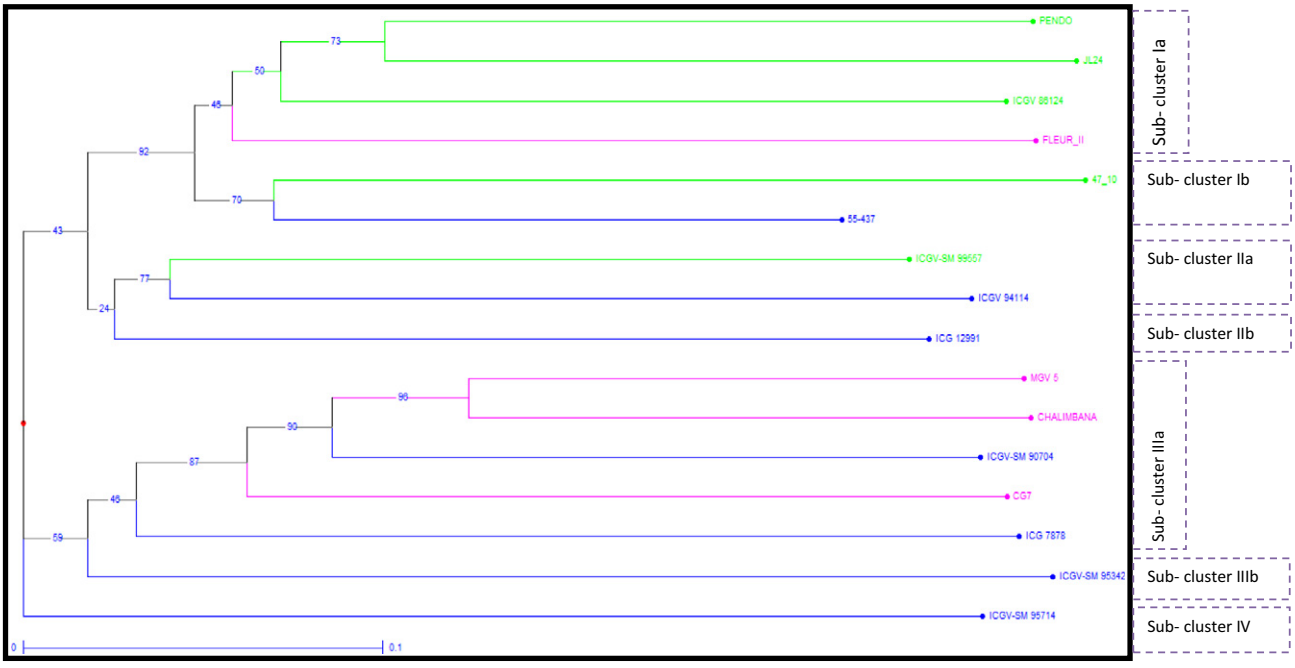

Fig. 2. Neighbor-joining tree illustrating the sub-clusters representing the 16 *Arachis* genotype, represented according to its predominant characteristic of disease resistance (green), yield and quality (pink) and farmer preferred traits (blue).

Table 3
Polymorphic SSRs loci identified in this study that were previously mapped to *Arachis* linkage groups (LG) (Gautami et al. [45], Wang et al. [30]).

LG	Markers									
a04 (LG9)	GM1062	Ap40	GM890	GM2246	TC11B04	GM1720	IPAHM105	GM2589	GM1919	GM1311
a09 (LG18)	GM2450	GM849	GM2359	GM1291	GM1911	PM675	AHGS0695	Ah1TC5D06	Ah1TC1D02	AHGS0993
a06 (LG5,10)	IPAHM659	GM1489	GM1490	GM2337	IPAHM245	TC11A04	GM1573	IPAHM689	GM1916	Ah2TC7C06
a03 (LG7)	GM1717	GM2402	GM2215	GM2528	GM2206	GM1954	Ah1TC0A01	pPGSseq19G7	AHGS0132	
a05 (LG19)	GM1049	GA34	GM1577	GM2078	RN16F05	GM1702	pPGSseq10D4	Ah1TC6E01	GA32	
b07 LG2)	GM1953	GM2156	GM2067	GM2073	GA24	GM2557	pPGPseq5D5	pPGSseq15C10		
a07 (LG4)	GM1494	GM1937	GM1076	GM1880	GM1986	GM1922	GM1990			
a08 (LG12)	GM2289	GM1628	GM2089	Ah1TC3B04	Ah2TC7A02	GM1713	GM2571			
b03 (LG14)	GM1854	GM1618	GM1996	GM2388	GM2009	Ah2TC9B12	GM2574			
b05 (LG21)	GM2137	GM1555	IPAHM136	GM1843	Ah1TC5D01	AHGS0729				
b01 (LG6)	GM1501	GM1331	Ah3	GM2607	pPGSseq13A7	AHGS0138				
b10 (LG5)	TC3E05	GM1742	GM2165	GM2032	Ah1TC1B02	Ah2TC11A02				
a10 (LG1)	GM2531	GM1788	GM2411	GA161	GM799					
b02 (LG16)	GM2196	Ah26	GA166	Ah1TC4F12						
b04 (LG13)	GM2584	GM1445	GM2033	AHGS0230						
b08 (LG4)	GM1961	IPAHM123	IPAHM606	GM1798						
(LG3)	GM2063	AHGS0369	AHGS0798	AHGS0278						
(LG17)	GM1821	pPGPseq2F5	GM1985							
(LG20)	AHGS0147	Ah2TC9H08	AHGS0151							
(LG11)	AHGS0357	pPGPseq1B9	GM1598							
b09	GM1483	Lec1								
(LG15)	GA166	Ah1TC4F12								
a02	RI1F06									

2. Materials and methods

2.1. DNA extraction

A total of 799 SSRs (supplementary data), comprising of di-and tri-nucleotide motifs from both genomic and expressed sequence tag (EST) SSRs, as compiled by Zhao et al. [32], were screened across 16 cultivated groundnut varieties indigenous to Africa. These varieties are listed in Table 1 and varied in yield and quality traits and tolerance to biotic stresses such as rust resistance (ICGV-SM 95342 and ICGV 94114), aphid resistance of GRD (ICG 12991) and virus resistance of GRD (ICGV-SM 90704), ELS resistance (ICGV-SM 95714 and ICG 7878), aflatoxin tolerance (55–437) high yield and quality traits (Fleur II, CG7/MGV4, MGV5 and Chalimbana), and other farmer preferred varieties (FPVs) (ICGV-SM 99557, Pendo, ICGV 86124, 47–10 and JL24/Luena).

Genomic DNA was extracted from 14-day old seedlings with one leaf from three individual plants combined into a single sample for each genotype. The genomic DNA was extracted according to the CTAB method of Mace et al. [33] with the exclusion of the phenol-chloroform extraction step.

2.2. SSR analysis

DNA from each variety were analyzed by PCR at the 799 selected SSR loci [32]. All forward primers contained an M13-tag (5′-CACGACGTTG TAAAACGAC-3′) on the 5′ end that was fluorescently labeled to allow detection of amplification products [34]. PCR amplification was performed in 10 μL and each reaction comprised of 1× PCR Buffer (20 mM Tris–HCl, pH 7.6; 100 mM KCl; 0.1 mM EDTA; 1 mM DTT; 0.5% Triton X-100; 50% glycerol), 2 mM $MgCl_2$, 0.16 mM dNTPs, 0.04 μM forward primer, 0.2 μM reverse primer, 0.16 μM fluorescent labeled M-13 tagged forward primer (FAM, VIC, NED PET), 0.2 U *Taq* DNA polymerase (SibEnzyme Ltd, Russia) and 30 ng DNA. PCR conditions were 94°C for 5 min, 35 cycles of 0.5 min at 94°C, 1 min at 59°C and 2 min 72°C and final extension at 72°C for 20 min using a GeneAmp® 9700 (Applied Biosystems). Amplification was confirmed by electrophoresis of PCR products (4 μL) on a 2% agarose gel against a 100 bp ladder (Fermentas), followed by capillary electrophoresis (ABI 3500 Genetic Analyzer) of successful PCR products. These (1.5–3.5 μL each) were co-loaded in sets of 4 markers together with the internal size standard, GeneScan™-500 LIZ® (Applied Biosystems). Gene Mapper Software (Version 4.0, Applied Biosystems) was used for allele scoring, followed by data analysis using PowerMarker Version 3.25 [35]. A dissimilarity matrix was compiled with DARwin software V5 [36].

3. Results and discussion

3.1. SSR marker properties and performance

A total of 799 markers (Supplementary data) were screened to identify the most informative markers for QTL mapping and pre/post-breeding applications.

Marker allele profiles obtained after capillary electrophoresis using GeneMapper 4.0, were graded on a scale of 1 to 4 for ease of scoring as illustrated in Fig. 1 (1 = clear single peaks, 2 = clear peaks with multiple stutter peaks, 3 = peaks not well defined but could be scored and, 4 = difficult to score due to noise, multiple loci binding or low availability). For grades 1, 2 and 3 the numbers of polymorphic markers obtained were 182, 61 and 133, respectively. In total, 423 markers were excluded from the final data set. These included 93 that were scored as grade 4, 169 that failed to amplify PCR products in the majority of the 16 varieties (i.e. availability <0.38) and 161 monomorphic markers. This screening provided 376 high quality polymorphic markers that worked well (average success rate of 94.2%) across the 16 varieties.

PowerMarker results were compiled for allele number, major allele frequency, how well each marker worked (availability), heterozygosity and PIC (Supplementary data).

Markers that were highly heterozygous confounded data interpretation and were carefully considered to determine if they had amplified two loci and if so, were split into two sets of alleles denoted with (_1/2) to the marker name. If both sets of alleles were heterozygous and polymorphic, these markers were retained. If one set of alleles was homozygous, this allele was discarded. Markers that would have resulted in two homozygous loci were not split. The total number of retained split markers was 18 and thus resulted in 394 polymorphic loci from a total of 376 markers.

The PIC range observed (0.06 for Ah-671 to 0.86 for Ah1TC4F12) in this study was similar to that reported by Pandey et al. [37] (PIC range 0.10 to 0.89). The mean PIC value obtained in the current study was

Table 3 (*continued*)

IPAHM108 Ah-671	GM2313	AHGS0347	AHGS0134	pPGSseq18C5	GM2480	Ah1TC5A07	Ah2TC7G10

0.49, with values above 0.5 observed in 174 (44%) of the loci analyzed, which was high compared to the findings of Cuc et al. [39], where only 15.7% of SSR markers showed PIC values >0.5. A study by He et al. [19] gave a lower percentage (34%) of polymorphic markers than that shown in our study. The number of polymorphic markers identified in this study (376 or 47% of the total number screened) was also high compared to other studies in groundnuts, which ranged from 3 to 33% of the markers analyzed [19,20,38]. However, the values were comparable to those reported by Cuc et al. [39] (44% with mean PIC 0.46; PIC range 0.12 to 0.75) and Mace et al. [40] (PIC range 0.29 to 0.60; mean PIC 0.47) with variations ascribed to genotype differences. The polymorphic markers identified in this study are therefore highly informative.

Marker GM2009 had a PIC value of 0.67 and has been shown to be closely associated with the major QTL for Late Leaf Spot (LLS) [23]. The genetic similarity of LLS and ELS disease resistance mechanisms [9] further supports the significance of this marker for QTL analysis for ELS resistance. Markers IPAHM108_1/2 and IPAHM123_2 had PIC values of 0.69/0.72 and 0.73 across 5 and 6 alleles, respectively. These were similar to that from a previous study by Cuc et al. [39] in which IPAHM108/123 had PIC values of 0.62/0.75 across 3 and 4 alleles, respectively. Other polymorphic IPAHM*x* markers varied from those of Cuc et al. [39] in terms of both low (IPAHM659_2/105/136/177) and high polymorphisms (IPAHM689) whilst allele numbers were fairly consistent in comparison. These variations in marker characteristics could be due to the inherent genotypic constitution of the cultivars used but cannot be confirmed as there were no common genotypes between this study and that of Cuc et al. [39]. Other markers that had high PIC values in this study as well as in that of Varshney et al. [41] were Ah1TC1E01, Ah1TC4F12 and Ah1TC6E01 with PIC values of 0.60–0.90. These similarities across different studies further highlight their usefulness in the present study across globally cultivated *Arachis* spp. The polymorphic markers identified in this study may also be useful across other legume species in comparative genomics studies as was ascertained with polymorphic soybean derived EST-SSRs in the *Arachis* genome [26]. These markers produced an average of 3.7 alleles per locus, for a total of 1476 alleles. The number of alleles per marker ranged from 2 to 11 with a mean of 3.74. Both higher numbers of alleles ranging from 2 to 14 [2,19] and lower numbers ranging from 2 to 8 [39] have been reported by previous studies. The most polymorphic markers with PIC values >0.70, reported by Hildebrand et al. [42] had allele values ranging from 5 to 11. The most informative SSR markers in this study were Ah2TC7H11, Ah1TC3E02, Ah1TC4F12, GNB70, Ah2TC11H06, AHGS0798, pPGPseq3B5, Ah2TC9H09, Lec1, Ah1TC2G05, AHGS0965, GA161, TC04G02, Ah1TC3B04, TC11A04, TC3E05, TC05A06 and GNB18 and had allele numbers ranging from 8 to 11 and PIC values of 0.78 to 0.86 and were considered important to distinguish all the varieties for use in MAS and other diversity studies.

Major Allele Frequency (MAF) ranged from 0.18 to 0.97 with a mean of 0.58 and heterozygosity ranged from 0 to 0.38 with a mean of 0.20. Markers with MAF between 0.5 and 0.8 (181 polymorphic markers in this study) have been reported to contribute approximately equally to information in linkage disequilibrium studies and should be useful in QTL mapping [43].

3.2. Genetic relationships and marker map locations

3.2.1. Dissimilarity matrix pair wise comparisons across the sixteen Arachis sp. varieties

A dissimilarity matrix was calculated from the allelic data of the 376 polymorphic markers (Table 2) and values ranged from 0.34 for the closest related varieties MGV5 and ICGV-SM 90704 to 0.66 for the most distant varieties Chalimbana and 47–10, with a mean value of 0.51. The dissimilarity values obtained were high in comparison to genetic distance values of previous studies in *Arachis* sp. [27,44] and ranged from 0.091 to 0.288 and 0.083 to 0.117, respectively. Subsequently, the most appropriate combinations for the development of bi-parental mapping populations for disease tolerance/resistance QTL mapping were identified, selecting the most distantly related varieties with contrasting expression of the trait and dissimilarity values above 0.5. As such, for ELS and LLS QTL mapping, ICG 7878 can be combined with FPVs 47–10, JL 24 and ICGV 86124 (dissimilarity values of 0.607, 0.597, 0.582 respectively) as well as with high yielding and quality trait variety FLEUR II (dissimilarity value: 0.57). ELS resistant genotype ICGV-SM 95714 will also combine well with FPVs 47–10, JL 24 and ICG 7878 (dissimilarity values: 0.566, 0.567 and 0.567 respectively). JL24 may be further improved by crossing with other resistant varieties such as ICG 12991 (GRD *Aphis* sp. resistant), ICGV-SM 90704 (GRD virus resistant), ICGV-SM 95432 (LLS and rust resistant) and ICGV-SM 95714 (ELS resistant) with dissimilarity values of 0.532, 0.563, 0.567 and 0.567, respectively. The matrix also indicated good varieties to combine in order to pyramid ideal abiotic

and resistance traits. In this regard, ICGV-SM 95714 (ELS resistant) will combine well with rust resistant ICGV 94114 and ICGV-SM 95432 (dissimilarity values 0.543 and 0.548 respectively) and drought tolerant *Aphis* sp. resistant ICGV 12991 with rust resistant ICG 95432 (0.572) and ELS resistant ICGV-SM 95714 (0.532) varieties. Other varieties may also be considered for pair wise introgression of disease resistance, such as rust resistant genotype ICGV-SM 95432 with *A. flavus* resistant 55–437 or ICGV 12991 and ICGV-SM 90704 for GRD resistance.

Sixty-three percent of the dissimilarity values calculated ranged from 0.50–0.66 and resulted from 237 polymorphic markers that could differentiate all varieties for the various traits of yield, quality and disease resistance. Nineteen percent of these values were associated with recommended crosses for introgression of ELS resistance. The high number of markers used in this study therefore enhanced the potential for targeted introgression of multiple disease resistance, yield and quality traits into farmer preferred and commercial groundnut varieties.

3.2.2. Genetic tree analysis

A neighbor-joining tree, illustrating the relatedness among the varieties, is presented in Fig. 2. The 16 varieties were grouped into three large clusters and a single outlier, ICGV-SM 95714. The majority of FPVs (47–10, ICGV 86124, JL 24 and Pendo) were grouped together in cluster 1 with ICGV 86124, 47–10, JL 24 and Pendo forming a more closely related sub-group (sub-cluster 1a). This may be attributed to low levels of out crossing [13,14,15]. Seed exchange among small holder farmers, planting proximity of preferred varieties, farmer preference for specific varieties and collection of seed for this study from a common geographic location may also have influenced the overall composition and relatedness of the varieties over the years. ELS resistant varieties ICG 7878 and ICGV-SM 95714 were noticeably distant from the majority of the varieties and hence more useful for trait QTL mapping and introgression into the other 14 varieties. ICGV-SM 95714 showed the lowest score for PCR performance across the varieties (90.9%), which may have contributed to its independent clustering.

3.3. Marker map distribution

A total of 139 (37%) of the 376 markers that were found to be polymorphic in this study have been previously mapped [30,45] (Table 3) and the number of markers per linkage groups (LG) and chromosomes (aa and bb) ranged from 0 for LG b06 to 18 for LG9 of chromosome a04. On average, the mapped markers were distributed evenly across all LGs with the exception of LG b06 of chromosome bb. These can be used to identify markers linked to various resistances and quality trait QTLs and their locations on the genome. The 139 is an appreciable number of mapped polymorphic SSRs since other studies successfully constructed genetic maps from 144 SSRs [46], 175 SSRs [47], 181/188 SSRs [23] and 324 SSRs [24] on recombinant inbred line populations as well as with larger marker numbers — 895 for the tetraploid 328 genome [45] and 1724 for the diploid genome [48].

4. Conclusions

In this study, 376 highly informative SSR markers were identified from 799 that were screened. This allowed genetic diversity assessment of 16 African groundnut cultivars with a wide repertoire of disease resistance and farmer preferred traits and a dissimilarity 'tool' was constructed that provides guidance on which parental combinations to use for mapping population development. In addition, 139 of these markers have been previously mapped and can now be employed in Quantitative Trait Loci (QTL) mapping. The additional 237 informative markers identified can be used to improve the efficiency of introgression of resistance to multiple important biotic constraints into farmer-preferred varieties of Sub-Saharan Africa.

Financial support

This study was supported by USAID Feed-the-Future Programme (EEM-G-00-04-00013) under the "Improving groundnut farmer incomes and nutrition through innovation and technology enhancement" (I-FINITE) project.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.ejbt.2014.10.004.

References

[1] Holbrook CC, Stalker T. Peanut breeding and genetic resources. Plant Breed Rev 2003;22:297–356. http://dx.doi.org/10.1002/9780470650202.ch6.
[2] Pandey MK, Monyo E, Ozias-Akins P, Liang X, Guimarães P, Nigam SN, et al. Advances in *Arachis* genomics for peanut improvement. Biotechnol Adv 2012;30:631–51. http://dx.doi.org/10.1016/j.biotechadv.2011.11.001.
[3] Izge AU, Mohammed ZH, Goni A. Levels of variability in groundnut (*Arachis hypogaea* L.) to *Cercospora* leaf spot disease — Implication for selection. Afr J Agric Res 2007;2:182–6 [cited 24 March 2014]. Available from Internet: http://www.academicjournals.org/journal/AJAR/article-abstract/D56E26030772.
[4] FAOSTAT. Database of food and agriculture organisation of the United Nations 2012. 2014 [cited 2 July 2014]. Available from Internet: http://faostat.fao.org/site/567/DesktopDefault.aspx?PageID=567#ancor.
[5] Singh F, Oswalt DL. Genetics and breeding of groundnut. Skill development series no. 4. Human resource development program. Patancheru, India: ICRISAT; 1991 [cited 24 March 2014]. Available from Internet: http://oar.icrisat.org/id/eprint/2754.
[6] Wu F, Narrod C, Tiongco M, Liu Y. The health economics of aflatoxin: Global burden of disease. Working paper; February, 2011 [cited 19 April 2014]. Available from Internet: http://www.ifpri.org/sites/default/files/publications/aflacontrol_wp04.pdf.
[7] Royal SS, Barry JB, Frederick MS, Daniel LC. Influence of broad leaf weeds on chlorothalonil deposition, foliar disease and peanut (*Arachis hypogaea* L) yield. Weed Technol 1997;11:51–8 [cited 24 March 2014]. Available from Internet: http://www.jstor.org/stable/3988229.
[8] McKenzie KS, Sarr AB, Mayura K, Bailey RH, Miller DR, Rogers TD, et al. Oxidative degradation and detoxification of mycotoxins using a novel source of ozone. Food Chem Toxicol 1997;35:807–20. http://dx.doi.org/10.1016/S0278-6915(97)00052-5.
[9] McDonald D, Subrahmanyam P, Gibbons RW, Smith DH. Early and late leaf spots of groundnut. Information Bulletin, 21. India: International crops research institute for the semi-arid tropics; 1985 [cited 24 March 2014]. Available from Internet: http://oar.icrisat.org/id/eprint/821.
[10] Cantonwine EG, Culbreath AK, Stevenson KL. Characterization of early leaf spot suppression by strip tillage in peanut. Phytopathology 2007;97:187–94 http://dx.doi.org/10.1094/PHYTO-97-2-0187.
[11] Podile AR, Kishore KG. Biological control of peanut diseases. In: Gnanamanickam SS, editor. Biological control of crop diseases. New York: CRC Press; 2002. p. 131–60. http://dx.doi.org/10.1201/9780203910955.ch7.
[12] Kokalis-Burelle N, Backman PA, Rodríguez-Kábana R, Ploper LD. Potential for biological control of early leaf spot of peanut using *Bacillus cereus* and chitin as foliar amendments. Biol Control 1992;2:321–8 http://dx.doi.org/10.1016/1049-9644(92)90026-A.
[13] Smartt J. Genetic instability and outcrossing in the groundnut variety Mani Pintar. Nature 1960;186:1070–1. http://dx.doi.org/10.1038/1861070a0.
[14] Reddy LJ, Nigam SN, Reddy AGS. Natural outcrossing in groundnut and its implications in groundnut breeding. J Oilseeds Res 1993;10:99–104 [cited 24 March 2014]. Available from Internet: http://oar.icrisat.org/id/eprint/1987.
[15] Knauft DA, Chiyembekeza AJ, Gorbet DW. Possible reproductive factors contributing to outcrossing in peanut (*Arachis hypogaea* L.). Peanut Sci 1992;19:29–31 http://dx.doi.org/10.3146/i0095-3679-19-1-7.
[16] Stalker HT, Mozingo LG. Molecular markers of *Arachis* and marker-assisted selection. Peanut Sci 2001;28:117–23. http://dx.doi.org/10.3146/i0095-3679-28-2-13.
[17] Luo M, Dang P, Bausher MG, Holbrook CC, Lee RD, Lynch RE, et al. Identification of transcripts involved in resistance responses to leaf spot disease caused by *Cercosporidium personatum* in peanut (*Arachis hypogaea*). Phytopathology 2005;95:381–7. http://dx.doi.org/10.1094/PHYTO-95-0381.
[18] Semagn K, Bjørnstad A, Xu Y. The genetic dissection of quantitative traits in crops. Electron J Biotechnol 2010;13. http://dx.doi.org/10.2225/vol13-issue5-fulltext-14.
[19] He G, Meng R, Newman M, Gao G, Pittman RN, Prakash CS. Microsatellites as DNA markers in cultivated peanut (*Arachis hypogaea* L.). Plant Biol 2003;3:3 http://dx.doi.org/10.1186/1471-2229-3-3.
[20] Moretzsohn MC, Hopkins MS, Mitchell SE, Kresovich S, Valls JFM, Ferreira ME. Genetic diversity of peanut (*Arachis hypogaea* L.) and its wild relatives based on the analysis of hyper variable regions of the genome. Plant Biol 2004;4:11 http://dx.doi.org/10.1186/1471-2229-4-11.
[21] Moretzsohn MC, Leoli L, Proite K, Guimaraes PM, Leal-Bertioli SC, Gimenes MA, et al. A microsatellite-based, gene-rich linkage map for the AA genome of *Arachis*

(Fabaceae). Theor Appl Genet 2005;111:1060–71 http://dx.doi.org/10.1007/s00122-005-0028-x.
[22] Asif M, Rahman MU, Mirza JI, Zafar Y. Parentage confirmation of cotton hybrids using molecular markers. Pak J Bot 2009;41:695–701.
[23] Sujay V, Gowda MVC, Pandey MK, Bhat RS, Khedikar YP, Nadaf HL, et al. Quantitative trait locus analysis and construction of consensus genetic map for foliar disease resistance based on two recombinant inbred line populations in cultivated groundnut (*Arachis hypogaea* L.). Mol Breed 2011;30:773–88 http://dx.doi.org/10.1007/s11032-011-9661-z.
[24] Qin H, Feng S, Chen C, Guo Y, Knapp S, Culbreath A, et al. An integrated genetic linkage map of cultivated peanut (*Arachis hypogaea* L.) constructed from two RIL populations. Theor Appl Genet 2011;124:653–64 http://dx.doi.org/10.1007/s00122-011-1737-y.
[25] Mondal S, Hadapad AB, Hande PA, Badigannavar AM. Identification of quantitative trait loci for bruchid (*Caryedon serratus* Olivier) resistance components in cultivated groundnut (*Arachis hypogaea* L.). Mol Breed 2014;33:961–73 http://dx.doi.org/10.1007/s11032-013-0011-1.
[26] Hong YB, Chen XP, Liu HY, Zhou GY, Li SX, Wen SJ, et al. Development and utilization of orthologous SSR markers in *Arachis* through Soybean (*Glycine max*) EST. Acta Agron Sin 2010;36:410–21. http://dx.doi.org/10.3724/SP.J.1006.2010.00410.
[27] Koppolu R, Upadhyaya HD, Dwivedi SL, Hoisington DA, Varshney RK. Genetic relationships among seven sections of genus *Arachis* studied by using SSR markers. Plant Biol 2010;10:15. http://dx.doi.org/10.1186/1471-2229-10-15.
[28] Mondal S, Badigannavar AM, D'Souza SF. Development of genic molecular markers linked to a rust resistance gene in cultivated groundnut (*Arachis hypogaea* L.). Euphytica 2012;188:163–73. http://dx.doi.org/10.1007/s10681-011-0619-3.
[29] Leal-Bertioli SCM, José ACVF, Alves-Freitas DMT, Moretzsohn MC, Guimarães PM, Nielen S, et al. Identification of candidate genome regions controlling disease resistance in *Arachis*. Plant Biol 2009;9:112 http://dx.doi.org/10.1186/1471-2229-9-112.
[30] Wang H, Penmetsa RV, Yuan M, Gong L, Zhao Y, Guo B. Development and characterization of BAC-end sequence derived SSRs, and their incorporation into a new higher density genetic map for cultivated peanut (*Arachis hypogaea* L.). Plant Biol 2012;12:10. http://dx.doi.org/10.1186/1471-2229-12-10.
[31] Yuan ML, Gong L, Meng R, Li S, Dang P, Guo B, et al. Development of trinucleotide (GGC)n SSR markers in peanut (*Arachis hypogaea* L.). Electron J Biotechnol 2010; 13:6. http://dx.doi.org/10.2225/vol13-issue6-fulltext-6.
[32] Zhao Y, Prakash CS, He G. Characterization and compilation of polymorphic simple sequence repeat (SSR) markers of peanut from public database. BMC Res Notes 2012;5:362. http://dx.doi.org/10.1186/1756-0500-5-362.
[33] Mace ES, Buhariwalla KK, Buhariwalla HK, Crouch JH. A high-throughput DNA extraction protocol for tropical molecular breeding programs. Plant Mol Biol Rep 2003;21:459–60. http://dx.doi.org/10.1007/BF02772596.
[34] Schuelke M. An economic method for the fluorescent labelling of PCR fragments. Nat Biotechnol 2000;18:233–4. http://dx.doi.org/10.1038/72708.
[35] Liu K, Muse SV. PowerMarker: An integrated analysis environment for genetic marker analysis. Bioinformatics 2005;21:2128–9 http://dx.doi.org/10.1093/bioinformatics/bti282.
[36] Perrier X, Jacquemoud-Collet JP. DARwin software. 2006 [cited 24 March 2014]. Available from Internet: http://darwin.cirad.fr/.
[37] Pandey MK, Gautami B, Jayakumar T, Sriswathi M, Upadhyaya HD, Gowda MVC, et al. Highly informative genic and genomic SSR markers to facilitate molecular breeding in cultivated groundnut (*Arachis hypogaea*). Plant Breed 2012;131: 139–47. http://dx.doi.org/10.1111/j.1439-0523.2011.01911.x.
[38] Song GQ, Li MJ, Xiao H, Wang XJ, Tang RH, Xia H, et al. EST sequencing and SSR marker development from cultivated peanut (*Arachis hypogaea* L.). Electron J Biotechnol 2010;13:3. http://dx.doi.org/10.2225/vol13-issue3-fulltext-10.
[39] Cuc LM, Mace ES, Crouch JH, Quang VD, Long TD, Varshney RK. Isolation and characterization of novel microsatellite markers and their application for diversity assessment in cultivated groundnut (*Arachis hypogaea* L.). BMC Plant Biol 2008;8: 55. http://dx.doi.org/10.1186/1471-2229-8-55.
[40] Mace ES, Phong DT, Upadhyaya HD, Chandra S, Crouch JH. SSR analysis of cultivated groundnut (*Arachis hypogaea* L.) germplasm resistant to rust and late leaf spot diseases. Euphytica 2006;152:317–30. http://dx.doi.org/10.1007/s10681-006-9218-0.
[41] Varshney RK, Mahendar T, Aruna R, Nigam SN, Neelima K, Vadez V, et al. High level of natural variation in a groundnut (*Arachis hypogaea* L.) germplasm collection assayed by selected informative SSR markers. Plant Breed 2009;128:486–94. http://dx.doi.org/10.1111/j.1439-0523.2009.01638.x.
[42] Hildebrand CE, Torney DC, Wagner RP. Informativeness of polymorphic DNA markers. Los Alamos Sci 1992;20:100–2.
[43] Goddard KAB, Hopkins PJ, Hall JM, Witte JS. Linkage disequilibrium and allele-frequency distributions for 114 Single-nucleotide Polymorphisms in five populations. Am J Hum Genet 2000;66:216–34. http://dx.doi.org/10.1086/302727.
[44] Tshilenge-Lukanda L, Nkongolo KKC, Kalonji-Mbuyi A, Kizungu RV. Epidemiology of the groundnut (*Arachis hypogaea* L.) leaf spot disease: Genetic analysis and developmental cycles. Am J Plant Sci 2012;3:582–8. http://dx.doi.org/10.4236/ajps.2012.35070.
[45] Gautami B, Fonceka D, Pandey MK, Moretzsohn MC, Sujay V, Qin H, et al. An international reference consensus genetic map with 897 marker loci based on 11 mapping populations for tetraploid groundnut (*Arachis hypogaea* L.). PLoS One 2012;7:e41213. http://dx.doi.org/10.1371/journal.pone.0041213.
[46] Varshney RK, Bertioli DJ, Moretzsohn MC, Vadez V, Krishnamurthy L, Aruna R, et al. The first SSR-based genetic linkage map for cultivated groundnut (*Arachis hypogaea* L.). Theor Appl Genet 2009;118:729–39. http://dx.doi.org/10.1007/s00122-008-0933-x.
[47] Hong Y, Chen X, Liang X, Liu H, Zhou G, Li S, et al. A SSR-based composite genetic linkage map for the cultivated peanut (*Arachis hypogaea* L.) genome. Plant Biol 2010;10:17. http://dx.doi.org/10.1186/1471-2229-10-17.
[48] Nagy ED, Guo Y, Tang S, Bowers JE, Okashah RA, Taylor CA, et al. A high-density genetic map of *Arachis duranensis*, a diploid ancestor of cultivated peanut. BMC Genomics 2012;13:469. http://dx.doi.org/10.1186/1471-2164-13-469.